

SCORE Search Results Details for Application 10807746 and Search Result 20061218_115218_us-10-807-746-7.rag.

[Score Home](#)
[Page](#)

[Retrieve Application](#)
[List](#)

[SCORE System](#)
[Overview](#)

[SCORE](#)
[FAQ](#)

[Comments /](#)
[Suggestions](#)

This page gives you Search Results detail for the Application 10807746 and Search Result 20061218_115218_us-10-807-746-7.rag.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2006, 19:53:34 ; Search time 199 Seconds
(without alignments)
1295.829 Million cell updates/sec

Title: US-10-807-746-7

Perfect score: 2937

Sequence: 1 MLRLNLRFLSFLLCISQSV.....KMTFFGSLDFSTLYFIQEKH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A. Geneseq 8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2937	100.0	564	8	ADT05582	Adt05582 Haemophil
2	2937	100.0	564	8	ADT51367	Adt51367 Non-typea
3	1023.5	34.8	558	7	ABO61169	Abo61169 Klebsiell
4	1021	34.8	605	7	ADF04606	Adf04606 Bacterial
5	1010.5	34.4	558	6	ABM69129	Abm69129 Photorhab
6	764.5	26.0	549	6	ABU30402	Abu30402 Protein e

7	763.5	26.0	549	5	AAO17804	Aao17804	H influen
8	757	25.8	542	6	ABU19675	Abu19675	Protein e
9	756	25.7	540	8	ADT05677	Adt05677	Haemophil
10	748	25.5	542	6	ABU21963	Abu21963	Protein e
11	737	25.1	574	6	ABU22049	Abu22049	Protein e
12	736.5	25.1	537	7	ABO63268	Abo63268	Klebsiell
13	728	24.8	546	6	ABU32103	Abu32103	Protein e
14	728	24.8	551	7	ABO64532	Abo64532	Klebsiell
15	726	24.7	535	9	AED82141	Aed82141	Hyperimmu
16	726	24.7	535	9	AED83036	Aed83036	Hyperimmu
17	725	24.7	535	6	ABU15032	Abu15032	Protein e
18	725	24.7	535	9	ADZ77674	Adz77674	Escherich
19	725	24.7	535	10	AEE97853	Aee97853	Escherich
20	725	24.7	535	10	AEE97710	Aee97710	Escherich
21	725	24.7	535	10	AEF18284	Aef18284	Dipeptide
22	725	24.7	536	6	ABM70366	Abm70366	Phototrab
23	718	24.4	535	6	ABU47545	Abu47545	Protein e
24	716.5	24.4	541	6	ABU39871	Abu39871	Protein e
25	715	24.3	517	6	ABU45336	Abu45336	Protein e
26	712	24.2	532	6	ABU39012	Abu39012	Protein e
27	712	24.2	535	4	AAU38208	Aau38208	Salmonell
28	712	24.2	535	6	ABU27536	Abu27536	Protein e
29	708	24.1	547	6	ABU22013	Abu22013	Protein e
30	705	24.0	535	6	ABU50401	Abu50401	Protein e
31	705	24.0	535	6	ABU41262	Abu41262	Protein e
32	705	24.0	539	7	ADF06592	Adf06592	Bacterial
33	698	23.8	547	6	ABU41922	Abu41922	Protein e
34	685.5	23.3	537	4	AAU36431	Aau36431	Pseudomon
35	685.5	23.3	537	6	ABU38751	Abu38751	Protein e
36	685.5	23.3	621	7	ABO77668	Abo77668	Pseudomon
37	677	23.1	533	6	ABU38752	Abu38752	Protein e
38	676	23.0	549	7	ABO77581	Abo77581	Pseudomon
39	664.5	22.6	563	7	ABO77669	Abo77669	Pseudomon
40	649.5	22.1	555	7	ABO75310	Abo75310	Pseudomon
41	648.5	22.1	1898	4	ABG25514	Abg25514	Novel hum
42	633	21.6	575	7	ABO77582	Abo77582	Pseudomon
43	483.5	16.5	549	2	AAW98830	Aaw98830	H. pylori
44	483.5	16.5	549	4	AAU35720	Aau35720	Helicobac
45	483.5	16.5	549	6	ABU30753	Abu30753	Protein e

ALIGNMENTS

RESULT 1
 ADT05582
 ID ADT05582 standard; protein; 564 AA.
 XX
 AC ADT05582;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Haemophilus influenzae (NTHi) protein - SEQ ID 618.
 XX
 RW middle ear bacterial infection; nasopharynx bacterial infection.
 XX
 OS Haemophilus influenzae.
 XX
 PN WO2004078949-A2.
 XX
 PD 16-SEP-2004.
 XX
 PF 05-MAR-2004; 2004WO-US007001.
 XX
 PR 06-MAR-2003; 2003US-0453134P.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL INC.
 XX
 PI Bakaletz LO, Munson RS, Dyer DW;
 XX
 DR WPI; 2004-662422/64.
 DR N-PSDB; ADT05581.
 XX
 PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
 useful for treating or preventing NTHi bacterial infections of the middle
 ear and/or nasopharynx.

applicant

XX
PS Claim 3; SEQ ID NO 618; 88pp; English.
XX
CC The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of *Haemophilus influenzae* (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC amino acid sequence represents an NTHi protein of the invention.
XX
SQ Sequence 564 AA;

Query Match 100.0%; Score 2937; DB 8; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.2e-253;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MLRLNLRFLSFLLCISQSVLQAAPSVPFTLTENGLTYCTHASGFSFNPQTADAGTSMNV 60
Db      1 MLRLNLRFLSFLLCISQSVLQAAPSVPFTLTENGLTYCTHASGFSFNPQTADAGTSMNV 60

Qy      61 VTEQIYNKLFDIKNHSATLTPMLAQSYSISADGKEILLNLRHGVKPHQTPWFTPTTRDFNA 120
Db      61 VTEQIYNKLFDIKNHSATLTPMLAQSYSISADGKEILLNLRHGVKPHQTPWFTPTTRDFNA 120

Qy      121 EDVVSINRVLGHNTYLTPLAEANVTYSNPQYRVFHEQARKVRFPPYFDSIKLEKIKSVT 180
Db      121 EDVVSINRVLGHNTYLTPLAEANVTYSNPQYRVFHEQARKVRFPPYFDSIKLEKIKSVT 180

Qy      181 ALSPYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQVK 240
Db      181 ALSPYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQVK 240

Qy      241 DYVYNQYVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASYPEVSGIGLL 300
Db      241 DYVYNQYVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASYPEVSGIGLL 300

Qy      301 KNDDKHYYMQSTDGMNLAYLAFNFDKPLMRDHEIRAAISQSLNRARIISHSYHNTATVAN 360
Db      301 KNDDKHYYMQSTDGMNLAYLAFNFDKPLMRDHEIRAAISQSLNRARIISHSYHNTATVAN 360

Qy      361 NIIPEVSWASTVNTPEFEDYHPKIAKNKLADKNLLNLWVINEEQVYNPAPFKMAEMIK 420
Db      361 NIIPEVSWASTVNTPEFEDYHPKIAKNKLADKNLLNLWVINEEQVYNPAPFKMAEMIK 420

Qy      421 WDLAQAGVKVKRAVTRPFLTAQLNRQSENYDLISLWLAGNLDPDGFMRRPILSCGTNE 480
Db      421 WDLAQAGVKVKRAVTRPFLTAQLNRQSENYDLISLWLAGNLDPDGFMRRPILSCGTNE 480

Qy      481 LTNLSNWCNEEFQFMDRAITTSLSLRAKAYNEAQELVRLPEIPIANVKRILVANSR 540
Db      481 LTNLSNWCNEEFQFMDRAITTSLSLRAKAYNEAQELVRLPEIPIANVKRILVANSR 540

Qy      541 VKGVKMTPFGLSDFSTLYFIQEKH 564
Db      541 VKGVKMTPFGLSDFSTLYFIQEKH 564
    
```

RESULT 2
ADT51367
ID ADT51367 standard; protein; 564 AA.
XX
AC ADT51367;
XX
DT 30-DEC-2004 (first entry)
XX
DE Non-typeable *Haemophilus influenzae* strain protein #1.
XX
KW antibacterial; auditory; antiinflammatory; antiarthritic; gene therapy;
KW vaccine; diagnosis; NTHi bacterial infection; otitis media; pneumonia;
KW sinusitis; septic arthritis; meningitis.
XX
OS *Haemophilus influenzae*.
XX
FN W02004087749-A2.
XX
PD 14-OCT-2004.
XX

PF 24-MAR-2004; 2004WO-US009021.
 XX
 FR 27-MAR-2003; 2003US-0458234P.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL INC.
 XX
 PI Bakaletz LO, Munson RS;
 XX
 DR WPI; 2004-737663/72.
 DR N-PSDB; ADT51361.
 XX
 PX New nontypeable strain of Haemophilus influenzae (NTHi) genes and
 PT polypeptides for diagnosing, preventing or treating NTHi bacterial
 PT infections, such as otitis media, pneumonia, sinusitis, septic arthritis
 PT or meningitis.
 XX
 PS Claim 5; SEQ ID NO 7; 93pp; English.
 XX
 CC The invention relates to an isolated polynucleotide comprising any of the
 CC 7 fully defined sequences of 810-2979 bp given in the specification. The
 CC encoded polypeptide comprises any of the 7 fully defined sequences of 269
 CC -992 amino acids given in the specification. The composition and methods
 CC are useful for diagnosing, preventing or treating NTHi bacterial
 CC infections, such as otitis media, pneumonia, sinusitis, septic arthritis
 CC or meningitis. This sequence corresponds to a protein from Haemophilus
 CC influenzae used in the invention.
 XX
 SQ Sequence 564 AA;

Query Match 100.0%; Score 2937; DB 8; Length 564;
 Best Local Similarity 100.0%; Pred. No. 1.2e-253;
 Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLRLNRLRFLSLCISQSVLQAQSPVPTFLTENGLTYCTHASGFSFNQPDAGTSMNV 60
 Db 1 MLRLNRLRFLSLCISQSVLQAQSPVPTFLTENGLTYCTHASGFSFNQPDAGTSMNV 60

Qy 61 VTEQIYNKLFQIKNHSATLTPMLAQSYISADGKEILLNLRHGKVPFQTPWFTPTROFNA 120
 Db 61 VTEQIYNKLFQIKNHSATLTPMLAQSYISADGKEILLNLRHGKVPFQTPWFTPTROFNA 120

Qy 121 EDVVFISINRVLGHTNYLPTLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKIKSVT 180
 Db 121 EDVVFISINRVLGHTNYLPTLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKIKSVT 180

Qy 181 ALSPYQVKIELFAPDSSILSHLASQYAIIFSQYAYQLSADDNLAQLDTHPVGTGPYQVK 240
 Db 181 ALSPYQVKIELFAPDSSILSHLASQYAIIFSQYAYQLSADDNLAQLDTHPVGTGPYQVK 240

Qy 241 DYVYNQYVRLVRNENYWKKEAKIEHIIIVDLSDRSGRLVKFFNNECQIASYPEVSIQIGLL 300
 Db 241 DYVYNQYVRLVRNENYWKKEAKIEHIIIVDLSDRSGRLVKFFNNECQIASYPEVSIQIGLL 300

Qy 301 KNDKRYHYMQSTDGMNLAYLAFNFDKPLMRDHEIPAAISQSLNRRARIHSTYHNTATVAN 360
 Db 301 KNDKRYHYMQSTDGMNLAYLAFNFDKPLMRDHEIPAAISQSLNRRARIHSTYHNTATVAN 360

Qy 361 NIIPEVSWASTVNTPEFEFDYHPKIAKNLADKNLLNLWVINEEQVYNPAPFKMAEMIK 420
 Db 361 NIIPEVSWASTVNTPEFEFDYHPKIAKNLADKNLLNLWVINEEQVYNPAPFKMAEMIK 420

Qy 421 WDLAQAGVKVKVRAVTRPFLTAQLRNQSENVDILSGWLAGNLDGDFMRPILSCGTKNE 480
 Db 421 WDLAQAGVKVKVRAVTRPFLTAQLRNQSENVDILSGWLAGNLDGDFMRPILSCGTKNE 480

Qy 481 LTNLSNWCNEEFQDFMDRAITTHSLSSRAKAYNEAQELVLRLEPIIPIANVKRILVANSR 540
 Db 481 LTNLSNWCNEEFQDFMDRAITTHSLSSRAKAYNEAQELVLRLEPIIPIANVKRILVANSR 540

Qy 541 VRGVMKTPFGSLDFSTLYFIQEKH 564
 Db 541 VRGVMKTPFGSLDFSTLYFIQEKH 564

RESULT 3
 ABO61169

ID AB061169 standard; protein; 558 AA.
 XX
 AC AB061169;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seqid 7686.
 XX
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 FN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M;
 XX
 DR WPI; 2003-895346/82.
 DR N-PSDB; ACH94720.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 7686; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SQ Sequence 558 AA;

Query Match 34.8%; Score 1023.5; DB 7; Length 558;
 Best Local Similarity 38.8%; Pred. No. 3e-82;
 Matches 209; Conservative 106; Mismatches 194; Indels 29; Gaps 7;

Qy	23	AAPSPV--TFLTENGLTYCTHASGFSFNQTDAGTSMNVVTEQIYNKLPDKNHSSATLT	80
		: : : : : : : : : : :	
Db	32	AAPALPDRADIRDSGFVYCVSGQVNTFNPKVSSGLIVDTLAAQIYDRLLDVOPTVRLV	91
Qy	81	PHLAQSYISADGKEILLNLNRHGVKHFQTFWPTPTDFNAEDVVSINRVLGHNTPLT	140
		: : : : : : : : : : :	
Qy	92	PELAESWEVLONGATYRPHLRHVFPFQTAWFTPTDFNADVIPTFGRIFNRD-----	145
Qy	141	AEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKISVTALSPYQVKIELFAPDSSILS	200
		: : : : : : : : : : :	
Db	146	-----HPWHNV-----NGSSFFYFDSLQFADSVESVRKLDNQTVFRLKRPDASFLW	192
Qy	201	HLASQYAIIFSQEYAYQLSADDNLQLDTHFVGTGPYQVKDYVYNQYVRLVRNENYWKKE	260
		: : : : : : : : : : : : : : : : : : : :	
Db	193	HLATHYASITSAEYAAARLTQDDRQELDRQPVGTGPFQLSDYRSQYVRLQRHFGYWRGK	252
Qy	261	AKIEHIVDLSDRGSLVKKFFNNECQIASYPEVSSQIGLLKNDKHHYMQSTDGMNLAFL	320
		: : : : : : : : : : : : : : : : :	
Db	253	PLMPQVVVDLGGSGGTGRLSKLLTGECVDLWPAASQLTLR-DDPRLRLTLRFGNIAWL	311
Qy	321	AFNFDKPLMRDHEIRAAISQSINRARIHSIYHNTATVANNIIPESWASVTNTPPEFED	380
		: : : : : : : : : : : : : : : : : : :	
Db	312	AFNTAKPPLDNPEVRHALALAINNQRMQSIYGTGAETASMLPRASWAYDNDKITE--	369
Qy	381	YHPKIAKNKLAD--KNNLLNLWVINEEQVYNPAPFKMAEMIKWDLAQAGVKVKVRAVTR	437
		: : : : : : : : : : : : : : : : : : :	
Db	370	YNPQEARARLKALGLENLTCLKWVPTSSQAWNPSPLKTAELIQADMAQIGVKVIIVPVEG	429
Qy	438	PFLTAQLRNQSENYDLILSGWLGNLDPDGFMRPILSCGTKNELTLNLSNCWNEEFQFMD	497

```

      | : | : | : | | | | : | | | | | : | | : | | | | :
Db      430 RFQEARLMDMS--HDLTSLGWATSDNDPDSFFRPLLSCAAIASQTNFAHWCNREFDDVLQ 487
QY      498 RAITTSLSRAKAYNEAQELVLRLEPIPIANVKRILVANSRVKGVKMTPPGSLDFS 555
      : | : : | | | | | | | : | | : | : | : | : : | : | : | :
Db      488 KALLSQQLSRRMDAYEAQRTLARELPVLPLASSLRQAYRYDMKGLVLSFPGNSAFA 545

```

RESULT 4

ADF04606

ID ADF04606 standard; protein; 605 AA.

XX

AC ADF04606;

XX

DT 12-FEB-2004 (first entry)

XX

DE Bacterial polypeptide #719.

XX

KW Proteus mirabilis infection; bacterial infection; antibacterial;
immunostimulant.

XX

OS Proteus mirabilis.

XX

PN US6605709-B1.

XX

PD 12-AUG-2003.

XX

PF 05-APR-2000; 2000US-00543681.

XX

PR 09-APR-1999; 99US-0128706P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Breton GL;

XX

DR WPI; 2003-895291/82.

DR

N-PSDB; ADF00434.

XX

PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.

XX

PS Disclosure; SEQ ID NO 4891; 870pp; English.

XX

CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polypeptide of the invention.

XX

SQ Sequence 605 AA;

Query Match 34.8%; Score 1021; DB 7; Length 605;
 Best Local Similarity 37.0%; Pred. No. 5.8e-82;
 Matches 202; Conservative 118; Mismatches 198; Indels 28; Gaps 8;

```

QY      16 SQSVELQAAPSVPTFLTENGLTYCTHASGFSFNPQTADAGTSMNVVTEQIYNKLFDIRKH 75
      : : | | | | | | | | | | : | | | : | : : | : | | : :
Db      67 TKATLSVAQE-PTNIHQNGFVYCVDSVNTFNPQLSSGLIIDPLAAQLYDRLLDVDPY 125

QY      76 SATLTPMLAQSYSISADGKEILLNLRHGVKHPQTWPFTPTDRFDNAEDVVSINRVLGHNT 135
      : | : | : | : | | | | | | | | | | : | : | | : | :
Db      126 TYRLIPEIAARWESLNGATYRFYLRKNVSPQTTPWFTPTRKLTADDVIFSERMFSAAN- 184

QY      136 YLPTLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKIKSVTALSYPQVKIELFAPD 195
      | | : | : | : | : | : | : | : | : | : | : | : | :
Db      185 -----NP-----YNOVNGGKYPYFDSLSLADNIKSIKKLKGKTYFENKSPD 226

```

Qy 196 SSILSHLASQYAIIFSQEYAYQLSADDNLQALDTHPVGTGPYQVKDYVYQVRLVRNEN 255
 Db 227 ASFLMHLATHYAFILSEEYASNLKSGNQSQLDWKPVGSGPFYLDQPGQGFVRLLRNQQ 286
 Qy 256 YWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASYPEVSIQGLKNDKHHYMQSTDGM 315
 Db 287 YWKGQPKMQQVVIDTGAGGTGRISKLLTGECVLAIPASQLKVLK-DDPRLRLTLRSGM 345
 Qy 316 NLAYLAFNFDKPLMRDHEIRAAISQSLNRARIHSIYHNTATVANNI IPEVSWASTVNT 375
 Db 346 NIAYLAFNNTKPPFNDLKVQAIAYAINNERLMGSIYYGTAETAASVLPASWAYD-NRA 404
 Qy 376 EPEFDYHPKIAKNKAD---KNLLNLWVINEEQVYNPAFPMKMAEMKWDLAQAGVKKV 432
 Db 405 KIT-EYNPEKSKQILKELGELKLNWVPSAPQSYNPSPLKMAELIQADLAQVGIGMNI 463
 Qy 433 RAVTRPFLTAQLRNQSENYDLILSGWLGNLDPDGFMRPILSCGTNKLNLNLCWNEEF 492
 Db 464 RPIEGRYQETSIMDR--HMTLSGWSTDSNDPSFTRPLFSCAIISSQTNLSHWCSPAF 521
 Qy 493 DQFMDRAITTHLSRRAKAYNEAQELVLRPLPIIANVKRILVANSRVKGVKMPGSL 552
 Db 522 DNVLQALYSQQLASRIDYHQADILAQELPVLPPLANSRMOQYRYDIKGLVLSTFGNA 581
 Qy 553 DFSTLY 558
 Db 582 SFAGVY 587

RESULT 5

ABM69129

ID ABM69129 standard; protein; 558 AA.

XX

AC ABM69129;

XX

DT 20-NOV-2003 (first entry)

XX

DE Photorhabdus luminescens protein sequence #2226.

XX

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.

XX

OS Photorhabdus luminescens.

XX

PN WO200294867-A2.

XX

PD 28-NOV-2002.

XX

PF 07-FEB-2002; 2002WO-IB003040.

XX

PR 07-FEB-2001; 2001FR-00001659.

XX

PA (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX

PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;

XX

DR WPI; 2003-148459/14.

XX

PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX

PS Claim 2; SEQ ID NO 2226; 1205pp; French.

XX

CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,

CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes produce vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi) that
CC are sensitive to P. luminescens-encoded toxins or antibiotics and
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens proteins

SX Sequence 558 AA;

Query Match 34.4%; Score 1010.5; DB 6; Length 558;
Best Local Similarity 36.3%; Pred. No. 4.5e-81;
Matches 205; Conservative 113; Mismatches 211; Indels 35; Gaps 6;

Qy	6	LRPLSLCLISQSVELQAA-----PSVPTFLTENGLTYCTHASGFSFNPQTADAGTS	57
Db	4	MSRLSYIWIILSLSPAIAETITTPKEKNHPVPTDIOQQGFITVCVNGNLNTPNPQLASSGLT	63
Qy	58	MNVVTEQYINKLFDIKNHSATLTPMLAQSYISADGKEILLNRHGKVFHQTPWFTPTRD	117
Db	64	VDTLAAQLYERLLDVPDYTYRLPELASHWEILONGATYRYLRHNVSFQSTDWFTPTRN	123
Qy	118	FNAEDVFISINRVLGHNITPLTAEANVTYNSNPQVRFEHQARKVRFPYFDSIKLNEKIK	177
Db	124	MAADVDVIFSFKRLFDKQHY-----YHNVNGHGYFPYFDSQLQADSITQ	164
Qy	178	SVTALSQYVKLELPADPSSNLSLSHAQVAILFSQYAYQLSADDNLAQLDTHPVGTGYPY	237
Db	165	SIRKLNKEYTVEFLNEPDASFLMHLATHYAPILSLQEQYQQQLHOMNRRHEQIWKVPVGTGF	224
Qy	238	QVKDYVYNQYVRLVRNENYWKKEAKIEHIVLDLSDRSGLKRVKFNNEQCIQASYEVSQY	297
Db	235	MLEDHRTQPIRLVLRHDKYWKGPQMKQIVLDVGAGGTGRMSKLTGECVCLAYAPASQL	284
Qy	298	GLLKNDDKHYYMSTDGMNLAYLAFNFKPLMRDHEIRAISQSLNRARIHSIYHNAT	357
Db	278	TVLR--DDPLRLTLRGRNMIYAFNTSKFPLDKIQVRAQYATNNQRMLQSIYGTAE	343
Qy	358	VANNIIEPVKSATVNTPEFEFDYHPKIAK---NKLADNLLNLNWINVEEQVYNPAPFK	414
Db	344	TASSILPRASWAYDNQTEITE---YNPEKSRRIINDLGLNLQLSLWVPSAQSYNPSPFK	401
Qy	415	MAEMIWDLAQAGVKVKVRAVTRPFLTAQLRNGSYNVDLILSGWLGNLDPGFMRPILS	474
Db	402	MAELIQADLAQVGITMSIKPVYGRFQETKIMDKS---HMDTLSGWSTDSDNDPSFRPLLS	459
Qy	475	CGTKNELTNSNWCNEEFDQMDRAITTSHLSSRAKAYNAEQVLVRELPIPIANVKRI	534
Db	460	CAAIASQTSFNSHWCPTFDKILREALINQQLLSRIKYHAAQQVLEQQLPILPLAYSLHL	519
Qy	535	LVANSRVGVKMTFPFGLSDPSTLY 558	
Db	520	QAYRHDIKGLVLSFPFGTSFAGVY 543	

```

RESULT 6
ABU30402
ID   ABU30402 standard; protein; 549 AA.
XX
AC   ABU30402;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #15929.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Haemophilus influenzae.
XX
PM   WO200277183-A2.
XX
YX

```


03-OCT-2002.
 21-MAR-2002; 2002WO-US009107.
 21-MAR-2001; 2001US-00815242.
 06-SEP-2001; 2001US-00948993.
 25-OCT-2001; 2001US-0342923P.
 08-FEB-2002; 2002US-00072851.
 06-MAR-2002; 2002US-0362699P.
 (ELIT-1) ELITRA PHARM INC.
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 WPI; 2003-029926/02.
 N-PSDB; ACA34272.
 New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.
 Claim 25; SEQ ID NO 58326; 1766pp; English.
 The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
[ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match 26.0%; Score 764.5; DB 6; Length 549;
Best Local Similarity 32.7%; Pred. No. 4.7e-59;
Matches 178; Conservative 100; Mismatches 227; Indels 39; Gaps 10;

Qy	16	SSQSVELQAAPSVPTFLTEGLTYITCTHASGFSFNPQADACTSMNVVTEQIYNKFLDIKNH	75
Db	24	SSSANKSTAQTEAKSSSNNTFVYCTAKAPLGFSPALIEGTSYNASSQSYNRLVEFKKG	83
Qy	76	SATITPMLAQSYSISADGKEILLNRHGWFKHOTPWTTTFDRFNADDDVFSINRVLGHNT	135
Db	84	STDIEPALAESWEISDDGLSYTEHLRGKGVFHTTKETPTTRFNADDDVFSFQRGLDPN	142
Qy	136	YLPTLAEANVTYSNPQYRVFHEQARKVRPYPFDSIKLNKKISVTSLSPYQVKIELFAPD	195
Db	143	-----HPYHNV-----SKGTYPYFKAMKPELLKSVKEVDONTIRITLNKTD	184
Qy	196	SSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGFGPYQVKDQVYNGYVLRVLN	255

Db	185	ATFLAAGMDPFIISYSAEYADSLMKAGKPETLDSRPVGTGPPFVVDYQTDQAIQYVAHEN	244
Qy	256	YWKKEAKRIEHIIVDLSTRSGVLVFNNECQIASYPEVSGIQLKLNDKHYHQYQSDGM	315
Db	245	YWRGRTPLDLRLVISIVPDATTRYAKLQAGTCDLILFENVADLAKMKTPKVQLLEQ-KGL	303
Qy	316	NLAYLAFNFDKPLMRDHIERAAISQSLNRARIHSIYHNATVANNIIPEVSWASTVNTPT	375
Db	304	NVAYIAFNTKEAFPQNVKQALNVAVDKAIIEAVYQAGTGSARKNLPPTII--SYNDE	361
Qy	376	EFEFDYHPKRIAKNKLD---KNLLNLNVLNIEEQVNPAPFFKMAEMIKWDLAQAGYKVK	431
Db	362	IQVYPYDPEKKAQLAEAGYPNGFETFDWIOPIVRASINPKRMKAEILIMADWAKIGVK--	419
Qy	434	VRAVTFP---FLTAQLRNQSENYDLI---LSGWLAGNLDPDGFMRPILSCGTKN-ELTNLSN	486
Db	420	---TNPTVYEWADYKRAKEGELTAGIFGWSDGNGDPNLFSPLL--GSSNIGNSNMAR	473
Qy	487	WCNEEFDQFMDRAITTHLSLRAKAYNEAQELVRLPIIPIANVKRILIVANSRVKGVKM	546
Db	474	FNNSEFDALINEAIGLTNKERAKLYKQAQVIVHNAQWPIVPAHVGFGFAPLSPRVKGYQV	533
Qy	547	TPFG 550	
Db	534	SPFG 537	

AAO17804

AA017804
ID AA017804 standard; protein; 549 AA.
XX
XX
AC AA017804;
XX
XX
DT 05-AUG-2002 (first entry)
XX
XX
DE H influenzae BVH-NTH13 protein SEQ ID NO: 6.
XX
KW Haemophilus influenzae infection; BVH-NTH11; otitis media; BVH-NTH12;
KW sinusitis; bronchitis; pneumonia; meningitis; bacteraemia; BVH-NTH13;
KW BVH-NTH14; BVH-NTH15; BVH-NTH16; BVH-NTH17; BVH-NTH18; BVH-NTH19;
KW BVH-NTH10; BVH-NTH11; BVH-NTH12; antiinflammatory; auditory;
XX anti-bacterial; vaccine.
XX
OS Haemophilus influenzae.
XX
PN WO200228889-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-CA001402.
XX
PR 02-OCT-2000; 2000US-0236712P.
XX
XX
FA (SHIR-) SHIRE BIOCHEM INC.
XX
PI Hamel J, Couture F, Brodeur BR, Martin D, Ouellet C, Tremblay M;
PI Charbonneau A, Vaysier C;
XX
DR WPI; 2002-435325/46.
DR N-PSDB; AAL46802.
XX
PT Novel isolated Haemophilus influenzae polypeptides BVH-NTH11-12, useful
PT for inducing protective immune responses against H. influenzae in animals
PT and for treating otitis media, sinusitis, bronchitis and pneumonia.
XX
PS Claim 17; Fig 6; 58pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC Haemophilus influenzae BVH-NTH11-12. The sequences can be used in the
CC production of a vaccine to protect against, and in the diagnosis of, H.
CC influenza infection, which can lead to otitis media, sinusitis,
CC bronchitis, pneumonia, meningitis and bacteraemia. The present sequence
CC is a protein of the invention
XX
XQ Sequence 549 AA.

Query Match	26.0%	Score 763.5;	DB 5;	Length 549;
Best Local Similarity	32.7%	Pred. No. 5.8e-59;		
Matches 178;	Conservative 100;	Mismatches 227;	Indels 39;	Gaps 10;
Qy	16	SQSVELQAAPSVPTFLTENGLTYCTHSGFSFNPQTADAGTSMNVVTEQIYNKLFIDKHN	75	
Db	24	SSSANKSTAQTEAKSSNNFTVYCTAKAPLGFSFALIEGTSSYNASSQQVYNRLVEFKKG	83	
Qy	76	SATLTPMLAQSYSISADGKEILLNLRHGVKFHOTPWFTPTRDFNAEDVVSINRVLGHNT	135	
Db	84	STDIEPALAESWEISDDGLSYTFHLRKGVKFMTTKEFTPTRDFNADDVVSFQRLQDPN	142	
Qy	136	YLPTLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKIKSVTALSPLYQVKIELFAPD	195	
Db	143	-----HPYHNV-----SKGTYPYFKAMKFPPELLKSVEKVDNTIRITINKTD	184	
Qy	196	SSILSHLASQYAIIFSQEQYAYQLSADDNLAQLDTHPVGTGPYQVKDYVYNQYVRLVRNEN	255	
Db	185	ATFLASLGMDFSIYSAEYADSMKAGKGPETLDSRPVGTGPVFVDYKTDQAQYVAHEN	244	
Qy	256	YWKKEAKIEHIIVDLSTDRSGRLVKFFNNBECQIASYPEVSQIGLLKNDKHHYMQSTDGM	315	
Db	245	YWKGRTPDLRLVISIVPDATTYAKLQAGTCDLILFPNVADLAKMKTDPKQVLEQ-KGL	303	
Qy	316	NLAYLAFNFDKPLMRDHEIRARISQSLNRARIHSIYHNTATVANNIIEPVSWASTVNTPT	375	
Db	304	NVAYIAFNTEKAPFNVKVRQALNYAVDKKATIEAVYQAGTSAKNPLPTTIW--SYNDE	361	
Qy	376	EFEFDYHPKIAKNKLAD---KNLLNLNVINEEQVYNPAPFKMAEMIKWDLAQAGVKVK	431	
Db	362	TQDYPYDPEKAKQLLAEAGYPNGFETDFWIQPIIRASNPMPKRMALIMADWAKIGVK--	419	
Qy	432	VRAVTRP--FLTAQLNRQNSYENLDI--LSGLWLAGNLDPDGMFMRPILSCGTRN-ELTNLSN	486	
Db	420	---TNPTVTEYAWDYKRKRAKELGTAGIFGWSGDNGDPDNFLSPLL--GSSNIGNSNMAR	473	
Qy	487	WCNEEFDFDMRAITTSLSHRAKAYNEAQELVIRELPPIIANVKRILVANSRVKGVKM	546	
Db	474	FNNSEFDALLNEAIGLTNKEERAKLYKQAVIVHQAQWIPVAHSVGFAPLSPRVKGVQY	533	
Qy	547	TPFG 550		
Db	534	SPFG 537		

RESULT 8

ABU19675

ID ABU19675 standard; protein; 542 AA.

XX

AC ABU19675;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #5202.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Borrelia cepacia.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

XX

PR 06-SEP-2001; 2001US-00948993.

XX

PR 25-OCT-2001; 2001US-0342923P.

XX

PR 08-FEB-2002; 2002US-00072851.

XX

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI; 2003-029926/02.

DR N-PSDB; ACA23545.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 25; SEQ ID NO 47599; 1766pp; English.

XX

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

- (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid;
- (2) a host cell containing the vector;
- (3) an isolated polypeptide whose expression is inhibited by the antisense nucleic acid;
- (4) an antibody capable of specifically binding the polypeptide;
- (5) producing the polypeptide;
- (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation;
- (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation;
- (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; or
- (9) obtaining an organism;
- (10) profiling a compound's activity;
- (11) a culture comprising strains in which the gene product is overexpressed or underexpressed;
- (12) determining the extent to which each of the strains is present in a culture or collection of strains; or
- (13) identifying the target of a compound that inhibits the proliferation of an organism.

The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *S. flexneri*, *S. flexneria*, or *E. coli*. These sequences are encoded by one of the target prokaryotic essential genes. The above sequence categories are patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX

SQ Sequence 542 AA;

Query Match 25.8%; Score 757; DB 6; Length 542;

Best Local Similarity 31.9%; Pred. No. 2.2e-58;

Matches 175; Conservative 88; Mismatches 242; Indels 44; Gaps 7;

Qy 18 SVELQAAPSVPTFLTENGLTYCTHASGFSEFNPQTADAGTSMNVVTEQIYNKLEFIKNHSA 77

: | | : | | : | | : | | : |

Db 20 AASLGVAGSAFAQIPNKTLVYCSEGSFAGFDSAQFTTGVDFTAATFTVYNRLVEFERGGT 79

Qy 78 TLTPMLAQSYSISADGKEILLNLRHGVKFHQTPWFTPTRDFNAEDVVFSINRVLGHNTYL 137

: | ||: : :|:| | :||| ||| | :| |||:| |:| ||| | | :| |

Db 80 KVEPGLAEKWDVSSDGKVYTFHLRHGVKFEHTTDFFKPTREFNADDVVFSEQRMLDPNNAF 139

Qy 138 PTLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEIKSVTALSPYQVKIELFAPDSS 197

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Db 140 -----RKAYPVSFYFTDMGLDKLITKVEKVDPYTVKFTLAEPNAP 180

[illegible]

QY 198 ILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQVKDYVYNQYVRLVRNENYW 257

Dk 181 ETONMAMEEACILGAEVCDLMKVGCDIADINCKDICTGRTTFRQUTRATETRRQRNKK 240

DD 161 FIQNMMEFASILSAEYGDQLMKAGRAADINQRPVGTGPFIFRSYTRDATIFRDGNPDYW 240

On 258 KE-EAKIEHTIIVDI STDRSCLVKEFENNECOTASYBEUSOICLI ENDDKHVYMOSTDCMI 316

Qy 250 RR-EARTHERIVDLSIDKSGRLVRKFNNNECQIASIFEVSSQIGLLRNDDKHFFIMQSFDGMN 316
||| |: |:: | : |||: ||| : ||| : ||| :

Db 241 KKGEVKLSKLIESTRDPGVVOKIKRNECOVMSYRPRADIATIK-ADSNVDMRISOACEN 288

241 ARGEVILDSKRIEFSITFDGVRVQRIRANECQVMSIFRFADIAILK-ADSNVDMFSQAGFN 299

Ov 317 LAYLAFNEDKPLMRDHEIRAAISOSINRARIHSHIYHNTATVANNIIPVSWASTVNTPE 376

[illegible]

Db 300 LGYLAYNVEHKPVVDKLEVROALDMAINKKAILESVYOGAGOAA SAPMPPTWS----- 352

[Home](#)
[About Us](#)
[Contact Us](#)
[Privacy Policy](#)
[Terms of Service](#)
[FAQ](#)
[Blog](#)
[Partners](#)
[Press](#)
[Careers](#)
[Feedback](#)

Qy 377 FEFDYHPKIAKNKLADKNLL-----NLWVINEEQVYNPAPFKMAEMIKWDLAQ 425

: | : | : | | || || : : : ||| ||||| : | | :

Db 353 --YDKNLKMAAYDTAKAKALLAKAGFPNGFEITLWAMPVQRAYNPARNLMAEMIQADWAK 410

Qy 426 AGVKVKVRAVTRPFLTAQLRNQSENYDLILSGWLAGNLDPDGFMRPILSCGTKNELTNLS 485

```

Db      411 IGVRKAKI--VTYEWGEYIKRAHAGEQDTMLIGWTGDNGDPDNLGLTGLGCEAIKG--NNFS 467
Qy      486 NMCNEEFDFQMDPRAITTSLSRAKAYNEAQLVRLPEIPIANVKRILVANSRVKGVK 545
       :|| : ||: : : ||: :| | | :||: : || |||| | :|
Db      468 HWCYKPFDELQVQKGRITTGQDARTKLYTQAQIFAQQLPFSPIANSTVYQPVKRNVDVDR 527
Qy      546 MTPFGSLDF 554
       : | | |
Db      528 IEPLGYARF 536

```

RESULT 9

ADT05677

ID ADT05677 standard; protein; 540 AA.

XX

AC ADT05677;

XX

DT 02-DEC-2004 (first entry)

XX

DE Haemophilus influenzae (NTHi) protein - SEQ ID 713.

XX

RW middle ear bacterial infection; nasopharynx bacterial infection.

XX

OS Haemophilus influenzae.

XX

PN WO2004078949-A2.

XX

PD 16-SEP-2004.

XX

PF 05-MAR-2004; 2004WO-US007001.

XX

PR 06-MAR-2003; 2003US-0453134P.

XX

PA (CHIL-) CHILDRENS HOSPITAL INC.

XX

PI Bakaletz LO, Munson RS, Dyer DW;

XX

DR WPI; 2004-662422/64.

XX

DR N-PSDB; ADT05676.

XX

PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.

XX

PS Claim 3; SEQ ID NO 713; 88pp; English.

XX

CC The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC amino acid sequence represents an NTHi protein of the invention.

XX

SQ Sequence 540 AA;

Query Match 25.7%; Score 756; DB 8; Length 540;

Best Local Similarity 32.4%; Pred. No. 2.6e-58;

Matches 180; Conservative 103; Mismatches 232; Indels 40; Gaps 11;

```

Qy      6 LRFLSFLLCIS-QSVELQAAPSVPTFLTENGLTYCTHASGFSFNQPTADAGTSMNVVTEQ 64
       :| : | : | : | : | :| :| :| :| :| :| :| :|
Db      4 LQLLWQLVQLINLASANKSTAQTEAKSSNNFTVYCTAKAPLGFSPALIIETGTSYNASSQQ 63
Qy      65 IYNKLFDIKNHSATLTPLMAQSYISADGKEILLNLRHGKVFHQTPTWFTPTTRDFNAEDVV 124
       :||:| : | | : | ||:| :| || :|| |||| | ||||| :|||
Db      64 VYNRLEVFKEKGSTDIEPALAESWEISDDGLSYTFHLRKGKVFHTTKEFTPTTRDFNADDVV 123
Qy      125 FSINRVLGHNTYLPDLAEANVTYSNPQYRVFHEQARKVRFYPFDSIKLNEKIKSVTALSP 184
       || | | | : | : | :| :| :| :| :| :| :| :| :| :|
Db      124 FSPORQLDPN-----HPYHNV-----SKGTYPYFKAMKFPPELLKRSVERVDD 164
Qy      185 YQVKIELFAPDSSILSHASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQVKDYVY 244
       :|| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      165 NTRITLNKTDATFLASLGMDFTSISYAEYADSLMKGKGPETLDSRFVGTGPFVFDYKT 224
Qy      245 NQVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASIYPEVSQIGLLKND 304

```

[illegible]

ABU21963

XX

DT 19-JUN-2003 (first entry)

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006

SCORE Search Results Details for Application 10807746 and Search Result 20061218_115220_us-10-807-746-7.rup.

[Score Home](#) [Retrieve Application](#) [SCORE System](#) [SCORE](#) [Comments /](#)
[Page](#) [List](#) [Overview](#) [FAQ](#) [Suggestions](#)

This page gives you Search Results detail for the Application 10807746 and Search Result 20061218_115220_us-10-807-746-7.rup.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2006, 19:54:23 ; Search time 304 Seconds
(without alignments)
1716.147 Million cell updates/sec

Title: US-10-807-746-7
Perfect score: 2937
Sequence: 1 MLRLNLRFLSFLLCISQSVE.....KMTFPFGSLDFSTLYFIQEKH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_7.2*
1: uniprot_sprot:*
2: uniprot_tramb1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2937	100.0	564	2	Q4QL73 HAEI8
2	2895.5	98.6	565	1	SAPA HAEIN
3	2219	75.6	540	2	Q714U3 HAEIN
4	1965	66.9	563	2	Q9CMC1_PASMU
5	1892.5	64.4	567	2	Q65U97_MANSM
6	1752.5	59.7	561	2	Q3EG24_ACTSC
7	1267.5	43.2	560	2	Q7VM01_HAEDU
8	1030.5	35.1	547	2	Q32FZ3_SHIDS
9	1027.5	35.0	547	2	Q83RL7_SHIFL
10	1025.5	34.9	547	2	Q8CW41_ECOL6
11	1023.5	34.8	547	1	SAPA_ECOLI
12	1023.5	34.8	547	2	Q7UCQ5_SHIFL
13	1022.5	34.8	547	2	Q31ZZ0_SHIDS
14	1022.5	34.8	547	2	Q3Z142_SHISS
					Q4q173 haemophilus
					P45285 haemophilus
					Q714u3 haemophilus
					Q9cmcl1 pasteurella
					Q65u97 mannheimia
					Q3eg24 actinobacil
					Q7vm01 haemophilus
					Q32fz3 shigella dy
					Q83rl7 shigella fl
					Q8cw41 escherichia
					Q47622 escherichia
					Q7ucq5 shigella fl
					Q31zz0 shigella bo
					Q3z142 shigella so

15	1022.5	34.8	549	1	SAPA_SALTY	P36634	salmonella
16	1022.5	34.8	549	2	Q5PCZ2_SALPA	Q5pcz2	salmonella
17	1022.5	34.8	557	2	Q57WX0_SALCH	Q57wx0	salmonella
18	1019.5	34.7	549	2	Q8Z7B5_SALTI	Q8z7b5	salmonella
19	1018.5	34.7	547	2	Q8X7F3_ECOS7	Q8x7f3	escherichia
20	1017.5	34.6	547	2	Q3MSG2_KLEOX	Q3msg2	klebsiella
21	1017.5	34.6	547	2	Q3MSI5_KLEPN	Q3msi5	klebsiella
22	1012	34.5	548	2	Q8ZE31_YERPE	Q8ze31	yersinia pe
23	1010.5	34.4	554	2	Q7N3X5_PHOLL	Q7n3x5	photorhabdu
24	1007.5	34.3	547	2	Q66A60_YERPS	Q66a60	yersinia ps
25	997.5	34.0	562	2	Q6D5R3_ERWCT	Q6d5r3	erwinia car
26	968	33.0	539	2	Q6LPF2_PHOPR	Q6lpf2	photobacter
27	964	32.8	540	2	Q8G187_ERWCH	Q8g187	erwinia chr
28	940.5	32.0	585	2	Q3NMU4_SODGL	Q3nmu4	sodalis glo
29	928	31.6	541	2	Q35U08_9GAMM	Q35u08	shewanella
30	915	31.2	541	2	Q2Z478_9GAMM	Q2z478	shewanella
31	914.5	31.1	535	2	Q3IF64_PSEHT	Q3if64	pseudotalter
32	910	31.0	541	2	Q366E0_9GAMM	Q366e0	shewanella
33	902.5	30.7	541	2	Q3NQ27_SHEFR	Q3nq27	shewanella
34	901.5	30.7	541	2	Q2WY77_9GAMM	Q2wy77	shewanella
35	899.5	30.6	541	2	Q2ZMV8_SHEPU	Q2zmv8	shewanella
36	898	30.6	542	2	Q3QDN5_9GAMM	Q3qdn5	shewanella
37	891	30.3	556	2	Q47XM8_COLF3	Q47xm8	colwellia p
38	886.5	30.2	539	2	Q87QH8_VIBPA	Q87qh8	vibrio para
39	883.5	30.1	541	2	Q3Q323_9GAMM	Q3q323	shewanella
40	882.5	30.0	525	2	Q8EC09_SHEON	Q8ec09	shewanella
41	879	29.9	547	2	Q33Q65_9GAMM	Q33q65	shewanella
42	877.5	29.9	541	2	Q5E0R7_VIBF1	Q5e0r7	vibrio fisc
43	871.5	29.7	541	2	Q8KUE4_VIBFI	Q8kue4	vibrio fisc
44	856	29.1	540	2	Q9KR62_VIBCH	Q9kr62	vibrio chol
45	854.5	29.1	541	2	Q5QUD5_IDILO	Q5qud5	idicmarina

ALIGNMENTS

RESULT 1

Q4QL73_HAEI8

ID Q4QL73_HAEI8 PRELIMINARY; PRT; 564 AA.

AC Q4QL73;

DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.

DT 19-JUL-2005, sequence version 1.

DT 07-FEB-2006, entry version 5.

DE ABC-type transport system, periplasmic component, involved in

DE antimicrobial peptide resistance.

GN Name=sapa; OrderedLocusNames=NTN11401;

OS Haemophilus influenzae (strain 86-028NP).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OX NCBI_TaxID=281310;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX PubMed:15968074; DOI=10.1128/JB.187.13.4627-4636.2005;

RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Munger R., Carson M.B.,

RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,

RA Munson R.S. Jr.;

RT "Genomic sequence of an otitis media isolate of nontypeable

RT Haemophilus influenzae: comparative study with H. influenzae serotype

RT d, strain KW20.;"

RL J. Bacteriol. 187:4627-4636 (2005).

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

DR EMBL; CP000057; AAX8824.1; -; Genomic DNA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000914; SBP_bac_5.

DR Pfam; PF00496; SBP_bac_5; 1.

DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.

KW Complete proteome.

SQ SEQUENCE 564 AA; 64420 MW; 0EB25C1FFA952643 CRC64;

Query Match 100.0%; Score 2937; DB 2; Length 564;

Best Local Similarity 100.0%; Pred. No. 1.3e-179;

Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLRLNLRFLSLLCISQSVLQAAPSVPFTLTENGLTYCTHASGFSFNPQADAGTSMNV 60
 Db 1 MLRLNLRFLSLLCISQSVLQAAPSVPFTLTENGLTYCTHASGFSFNPQADAGTSMNV 60

Qy 61 VTEQIYNKLFDIKNHSATLTPMLAQSYSISADGKEILLNLRHGKVFHQTFWFTPTDRFNA 120
 Db 61 VTEQIYNKLFDIKNHSATLTPMLAQSYSISADGKEILLNLRHGKVFHQTFWFTPTDRFNA 120

Qy 121 EDVVFSLNRVLGHNTYLPTEAEANVTYSNPQYRVFHEQARKVRFPPYFDSIKLEKIKSVT 180
 Db 121 EDVVFSLNRVLGHNTYLPTEAEANVTYSNPQYRVFHEQARKVRFPPYFDSIKLEKIKSVT 180

Qy 181 ALSPYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQVK 240
 Db 181 ALSPYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQVK 240

Qy 241 DYYVNYQVRLVRNENYWKKEAKIEHIIVDLSTDRSGRLVKFFNNECQIASYPEVSIQIGLL 300
 Db 241 DYYVNYQVRLVRNENYWKKEAKIEHIIVDLSTDRSGRLVKFFNNECQIASYPEVSIQIGLL 300

Qy 301 KNDDKHYYMQSTDGMNLAFLAFNFKPLMRDHEIRAAISQSLNRARIHSIYHNTATVAN 360
 Db 301 KNDDKHYYMQSTDGMNLAFLAFNFKPLMRDHEIRAAISQSLNRARIHSIYHNTATVAN 360

Qy 361 NITPEVSWASTVNTPEFEDYHPKIAKNKLADNLLNLWVINEEQVYNPAPFMAEMIK 420
 Db 361 NITPEVSWASTVNTPEFEDYHPKIAKNKLADNLLNLWVINEEQVYNPAPFMAEMIK 420

Qy 421 WDLAQAGVKVKVRAVTRPFLTAQLRNQSENVDLISGLWAGNLDPDGFMRFILSCGTKE 480
 Db 421 WDLAQAGVKVKVRAVTRPFLTAQLRNQSENVDLISGLWAGNLDPDGFMRFILSCGTKE 480

Qy 481 LTNLSNWCNEEFQFMDRAITTHLSSRAKAYNEAQELVRLRELPPIPIANVKRILVANSR 540
 Db 481 LTNLSNWCNEEFQFMDRAITTHLSSRAKAYNEAQELVRLRELPPIPIANVKRILVANSR 540


Qy 541 VKGVKMPFPGSLDFSTLYFIQEKH 564
 Db 541 VKGVKMPFPGSLDFSTLYFIQEKH 564

RESULT 2

SAPA_HAEIN

ID SAPA_HAEIN STANDARD; PRT; 565 AA.
 AC P45285;
 DT 01-NOV-1995, integrated into UniProtKB/Swiss-Prot.
 DT 01-NOV-1995, sequence version 1.
 DT 07-MAR-2006, entry version 37.
 DE Peptide transport periplasmic protein sapA precursor.
 GN Name=sapA; OrderedLocusNames=HI1638;
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OC NCBI_TaxID=727;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Rd / Kw20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -!- FUNCTION: Involved in a peptide intake transport system that plays
 CC a role in the resistance to antimicrobial peptides (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
 CC -!- SIMILARITY: Belongs to the bacterial solute-binding protein 5

CC family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL: L42023; AAC23285.1; -: Genomic_DNA.
 DR PIR: A64134; A64134.
 DR HSSP: P23847; 1DPE.
 DR GenomeReviews: L42023_GR; H11638.
 DR TIGR: H11638; -.
 DR BioCyc: HNW71421:H11638-MONOMER; -.
 DR InterPro: IPR000914; SBP_bac_5.
 DR Pfam: PF00496; SBP_bac_5; 1.
 DR PROSITE: PS01040; SBP_BACTERIAL_5; 1.
 KW Complete proteome; Peptide transport; Periplasmic; Protein transport;
 KW Signal; Transport.
 FT SIGNAL 1 23 Potential.
 FT CHAIN 24 565 Peptide transport periplasmic protein
 FT sapA.
 FT /FTId=PRO_0000031803.
 SQ SEQUENCE 565 AA; 64504 MW; 449E454F1278C2A7 CRC64;

Query Match 98.6%; Score 2895.5; DB 1; Length 565; 
 Best Local Similarity 98.6%; Pred. No. 5.9e-177;
 Matches 558; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Qy 1 MLRLNRLFLSLFLCISQSVELQAAPSVPFTLTENGLTYCTHSGFSFNPQTADAGTSMNV 60
 Db 1 MLRLNRLFLSLFLCIIQSVELQAAPSVPFTLTENGLTYCTHSGFSFNPQTADAGTSMNV 60
 Qy 61 VTEQIYNKLFIDIKNHSATLTPMLAQSYSISADGKEILLNLRHGVKPHQTPWFTPTRDFNA 120
 Db 61 VTEQIYNKLFIDIKNHSATLTPMLAQSYSISADGKEILLNLRHGVKPHQTPWFTPTRDFNA 120
 Qy 121 EDVVFSINRVL-GHNTYLPPLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKIKSV 179
 Db 121 EDVVFSINRVLGGHNTYLPPLAEANVTYSNPQYRVFHEQARKVRFPYFDSIKLNEKIKSV 180
 Qy 180 TALSPLYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQV 239
 Db 181 TALSPLYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLAQLDTHPVGTGPYQV 240
 Qy 240 KDYYVNYQYVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASYPEVSIQGL 299
 Db 241 KDYYVNYQYVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASYPEVSIQGL 300
 Qy 300 LKNDKHYMQSTDGMNLAYLAFNFDKPLMRDHEIRAAISQSLNRARIHISIHNTATVA 359
 Db 301 LKNDKHYMQSTDGMNLAYLAFNFDKPLMRDHEIRAAISQSLNRARIHISIHNTATVA 360
 Qy 360 NNIIPEVSWASTVNTPEFEFDYHPKIAKNLADKNLLNLWVINEEQVYNPAPFKIAEMI 419
 Db 361 NNIIPEVSWASTVNTPEFEFDYHPKIAKNLADKNLLNLWVINEEQVYNPAPFKIAEMI 420
 Qy 420 KWLDAQGVKVKVRAVTRPFLTAQLRNSQENYDLILSGWLAGNLDPPGFMRLILSCGTKN 479
 Db 421 KWLDAQGVKVKVRAVTRPFLTAQLRNSQENYDLILSGWLAGNLDPPGFMRLILSCGTKN 480
 Qy 480 ELTNLSNWCNEEFQDMRAITTSLSLRKAYNEAQELVRLRELPPIPIANVKRILVANS 539
 Db 481 ELTNLSNWCNEEFQDMRAITTSLSLRKAYNEAQELVRLRELPPIPIANVKRILVANS 540
 Qy 540 RVKGVMTPFGSLDFSTLYFIQEKH 544
 Db 541 RVKGVMTPFGSLDFSTLYFIQEKY 565

RESULT 3
 Q71403_HAEIN
 ID Q71403_HAEIN PRELIMINARY; PRT; 540 AA.
 AC Q71403;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 8.
 DE SapA.
 GN Name=sapA;

```

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OK NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=700222;
RX MEDLINE=22999298; PubMed=14638817;
RX DOI=10.1128/IAI.71.12.7202-7207.2003;
RA Satola S.W., Schirmer P.L., Farley M.M.;
RT "Genetic analysis of the capsule locus of Haemophilus influenzae
RL serotype f.";
RL Infect. Immun. 71:7202-7207(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF549211; AAC12665.1; ~; Genomic DNA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; F:transport; IEA.
DR InterPro: IPR0005914; SBP_bac.5.
DR Pfam: PF00496; SBP_bac.5.1.
SQ SEQUENCE 540 AA; 61739 MW; 4BB46B7411611B48 CRC64;

```

Query Match 75.6%; Score 2219; DB 2; Length 540;
Best Local Similarity 76.8%; Pred. No. 1.2e-133;
Matches 414; Conservative 62; Mismatches 61; Indels 2; Gaps 2;

QY	1	MLRLNRLFLSFLLCISQSVEL-QAAPSVPPTFLTENGTLTCYTHASGFSFNPQTADAGTSMN	59
Db	1	MLHRNVTF-CFLCGLGSLNLINLQAAPRIPKMLTENGTLTCYTHASGFSFNPQTADAGTSMN	59
QY	60	VVTEQIYNKRLFDIKNHSATLTPMLAQSYSISADGKEILLNLRHGKVFHQTWPFTPTRDFN	119
Db	60	VVTEQIYNKRLFDHKDSAAALVFLPAQSYSISSDGKILLNLRGQVFKHRTWPFTSTREFN	119
QY	120	AEDVVFSINRVLGHNTYLPFLASINRNVHTYLPNQVYRFLHGHQARKVRFPYFDSIKLNEKIRSV	179
Db	120	AEDVVFSINRVLGHDTYLPFLASINRNVHTYLPNQVYRFLHGHQARKVRFPYFDSIKLNEKIRSV	179
QY	180	TALSPQVKIELFAPDSISILSHLASQYAIIFSQEVAYQLSADDNLAQLDTHPVGTGPGYQV	239
Db	180	TATNPQVKIELFEPDASILSHLASQYSIIFSQEVAYQLSADDNLSGLDTHPVGTGPGYQV	239
QY	240	KDYVYNQVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVFNKNNECQIASYPEVSQIGL	299
Db	240	KDYVYNQVRLIRNEEYWKKEAKIKNIIVDLSEKSGRLIKFFNNECQIASPESQIGL	299
QY	300	LKNDKHYYMOSTEGNNLAYLAFNPFKPLMRDHEIRAAISQSNIRAKTIHSIYHNTATVA	359
Db	300	LSEKNASYYLQSTEGNNLAYLAFNFQKSLMQDITIRAIQSQSNIRFRIVRIYHNTATVA	359
QY	360	NNIIPESVSASTVNTPEFDFYHPKIAKKILADKNLLMLVINEEQVYNPAPFKMAEMI	419
Db	360	NNIIPDISWASAINTPDPTDYDQPSKAEILRDKKILAKMWVINEEQVYNPAPFKMAELI	419
QY	420	KWDLAQAGVKKVRVATRFPLTAQLNRQSENYYDLISGLWAGNLDPDGMFRLPSCGTGN	479
Db	420	KWDLAKVGVDKVRSVTRFLTEQLGNHEDYDLITGLWAGNLDPDGMFRLPSCDTGN	479
QY	480	ETLNLNNEEFPQDMFDRATITSHLSKKAYNAEGLVIRELPIDANVRKILVAN	538
Db	480	ETLNLNWCNPFEPKDMDRALSTNHLVERSKAYNSAEGILNELPIPIVIANVORILVAS	538

```

RESULT 4
Q9CMC1_PASMU
ID Q9CMC1_PASMU PRELIMINARY; PRT; 563 AA.
AC Q9CMC1;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE SapA.
GN Name=sapA; OrderedLocusNames=PMO911;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```

```

OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA Ray B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AE006130; AAK02995.1; -; Genomic_DNA.
DR HSSP; P23847; LDPE.
DR BioCyc; PM0U747:PM0911-MONOMER; -.
DR GO; GO:005215; F:transporter activity; IEA.
DR GO; GO:006810; P:transporter; IEA.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5_1.
DR Complete proteome.
SQ SEQUENCE 563 AA; 64532 MW; 9F143828AC2C8306 CRC64;

```

```

Query Match          66.9%; Score 1965; Db 2; Length 563;
Best Local Similarity 65.7%; Pred. No. 2.4e-117;
Matches 366; Conservative 89; Mismatches 100; Indels 2; Gaps 1;

Qy 1 MLRLNRLFLSLFLICISQSVELQAAPSVPFTLTENGITYCTHAGSFNFNQTDAGTSMNV 60
Db 1 MLIRKVI FACFLFLYSHFV--TAAPRVPNELTQNGLIYCTHATGFSNFNQTDAGTSMNV 58

Qy 61 VTEQIYNKLFDIKNHSATLTPMLAQSYSISADGKEILLNLRHGKVFHQTPWFTPTPRDNA 120
Db 59 ITEQIYNKLFTFSDSATVPSLAEYSRVSDNGTLTITNLRKGKVFHHTWETPTPRDNA 118

Qy 121 DDVFSINRLGVHNTYLTPLAEANVTYSNPQYRVFHEQARKRVFPFYDSIKLNEKIKSVT 180
Db 119 DDVFSINRLGLYSYLTPLDDES IHYSNPQYRVFHFQAKKRPFPYFESIKLNQKIKSIK 178

Qy 181 ALSPYQVRIELFAPDSSILSHLSAQYAIIEQYVAYLQADNLQALDTHTPVGTGPYQVK 240
Db 179 AITPYQVQIKLFQADASILSHLSAQYAIIEQYVAYLQADNLVQLDLTPVGTGPYQVK 238

Qy 241 DYVYNQYVRLVRNENYWKKEAKIEHIVDLSDRSGRLVKLFNNECQLASYPEVSQLGL 300
Db 239 NYFNNQYVRFIRNEHWKKPAQIKNIIDLSLSDRTGRVLKFLNGECQIYSPYEVSQLGL 298

Qy 301 KNDDEKHYMSTGDMNLAYLAFNDFPIMLRDHEIRRAISOSLNRRARIHSIYHNTATVAN 360
Db 299 QDKNEHFYDVFEGMNLSYLAFNFKPAMKSMKLRRAISQIDRHRITVQTYIYHNTATVAN 358

Qy 361 NIPEVSWASTVNPPEFFDFYHPKIAKKNLKADLNLLNLAVNINEEQVYNPAFFKMAEMIK 420
Db 359 NIPISISWASKVNTPDFAYDYAPEKARAFQDKQLQLTWVWINEEQVYNPSPLKMAELIK 418

Qy 421 WDLAQAGVKVKRAVTRFFLTIAQLRNQSENYDLISGLWAGNLDPDGMFMPILSCGTNE 480
Db 419 RDLANVGKVIQVPTVRTYILIERLKAHSEYDMLIAGLWAGNLDPDGMFMPILSCNTVE 478

Qy 481 LTNLNSNCNEEFQDFMDRAITTHLSLSRAKAYNAEQELVLRLELPIPIANVRKRLVANSR 540
Db 479 ITFNNSNCDFLDFHMDNNTNMLLRASEYNLAQELISEVPLPIANAKRMLVSPN 538

Qy 541 VKGVKMTNPGSLDFSTL 557
Db 539 VOGVKMSPPGSGINFNI 555

```

```

RESULT 5
Q65U97_MANSM
ID   Q65U97_MANSM      PRELIMINARY;      PRT;   567 AA.
AC   Q65U97;
DT   25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT   25-OCT-2004, sequence version 1.
DT   07-FEB-2006, entry version 10.
DE   OpaA protein.

```

GN Name: oppA; OrderedLocusNames=MS0856;
 OC Mannheimia succiniciproducens (strain MBEL55E).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OC NCBI_TaxID=221988;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15378067; DOI=10.1038/nbt1010;
 RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
 RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;
 RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
 RT succiniciproducens.";
 RL Nat. Biotechnol. 22:1275-1281(2004).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AE016827; AAU37463.1; -; Genomic DNA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000914; SBP_bac_5.
 DR Pfam; PF00496; SBP_bac_5; 1.
 KW Complete proteome.
 SQ SEQUENCE 567 AA; 64687 MW; 6DA1B590A970B46D CRC64;

Query Match 64.4%; Score 1892.5; DB 2; Length 567;
 Best Local Similarity 64.9%; Pred. No. 1.le-112;
 Matches 362; Conservative 84; Mismatches 103; Indels 9; Gaps 4;

Qy 8 FLFLLCISQSVELQ-AAPSVPTFLTENGTYCTHAGSFNFQCTADAGTSMNVTEQY 66
 Db 8 FIGFLFSAMLFFSWAAPRVPEILTQNGLIYCTHSSGFSNFQCTADAGTSMNVTEQY 67

Qy 67 NKLFDIKNHSATLTPMLAQSYSISADGKEILLNRHGVKHFQTPWFTPTROFNAEDVVF 126
 Db 68 NKLFEIKNNSRLFSLAQSYKISEDGKTTIVLRKGVFHTPTWFTPSRNFNAEDVVF 127

Qy 127 LNRVLGHNTLYPTLAEANVTYSNP----QYRVFHEQARKVRFPYFDSIKLNKIKSVTAL 182
 Db 128 LNRVLGHNTSLP---EFNASEQQQGHKRYNIFHELAKKTRFPYFDSIKLNKIKSVTAL 184

Qy 183 SPYQVKIELFAPDSSILSHLASQYAIIFSQEYAYQLSADDNLQALDTHPVGTGPYQVKDY 242
 Db 185 DPTVQINLFPADASILSHLASQYAIIFSHAYALQNLADDNLQALDLPVGTGPYQVKDY 244

Qy 243 VYNQYVRLVRNRYNWKKEAKIEHIVDLSTDRSGRLVKFFNNECQIASPEVSIQGLKN 302
 Db 245 FRNQYVRLIRHENYWKKEAIEKNIIIDLSPTDRGLAKFFNNECQIAAPDVSQGLGLQE 304

Qy 303 DDKHYMQSTOGMNLAYLAFNFXFLMRDHEIRAAISQSLNRIIHSIYHNTATVANNI 362
 Db 305 NGERFQTLSDGMNLAFLAFNFKRFLMQDAEIRGIAQAINRHRIKIDYNYNTASVANKI 364

Qy 363 IPEVSWA-STVNTPEFDFYHPKIAKKLADKNLLMLAVINEEQVYNPAPFMAEMIKW 421
 Db 365 IPEVSWAGSDSNHNSFAYDYDPAQAKKVLQDRQLSDMLVLEQVYNPSPKMAELIKH 424

Qy 422 DLAAQGVKVKVRAVTRFILTALRNQSENVDLILSGVLNGLDPDGFMPILSCGTKNEL 481
 Db 425 DLTKAGIEVKVRLIRSNFLMEQNRNSENVDLILGGLAVSLDPDSFMPILSCGTSTSEI 484

Qy 482 TNLNWCNEPEDFQFQWRAITTSLLSHSRKAYNEAQELVRLREPIPIANVKRILVANSRV 541
 Db 485 TNLNWCNSQSFEILDRALISNTNERAVNYHLAEQVLSLPIPIASVKRILISNSNV 544

Qy 542 KGVKMTFPGSLDFSTLYF 559
 Db 545 QGVEMSPFGSISFEKLSF 562

RESULT 6
 Q3EG24.ACTSC
 ID Q3EG24.ACTSC PRELIMINARY; PRT; 561 AA.
 AC Q3EG24;
 DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 08-NOV-2005, sequence version 1.

```

DT 07-FEB-2006, entry version 3.
DE ABC-type dipeptide/oligopeptide/nickel transport systems, periplasmic
DE components precursor.
GN ORFNames=AsucDRAFT_0543;
OC Actinobacillus succinogenes 1302.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=339671;
RN [1]
RF NUCLEOTIDE SEQUENCE.
RC STRAIN=1302;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Actinobacillus
RT succinogenes 1302.";
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RF NUCLEOTIDE SEQUENCE.
RC STRAIN=1302;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Actinobacillus
RT succinogenes 1302.";
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAKK01000026; EAO50273.1; -; Genomic DNA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000914; SBP_bac 5.
DR Pfam; PF00496; SBP_bac 5; 1.
DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.
KW Signal.
FT SIGNAL 1 21 Potential.
FT SIGNAL 561 561 Potential.
SQ SEQUENCE 561 AA; 63813 MW; 73AD17987DB0757 CRC64;

```

[illegible]

Db	414	EMIRTDINAVGLNVRKVLISNFLENMLNKETEDYDLILAGLIASSLDLDPDSFLRPLISCD	473
Qy	477	TKNELTNLSNWCNEEFQDFQMDRAITTHSLSSRAKAYNAQGLVLRPLPIIPIANVKRILV	536
Db	474	TTDEVSNVSNWCNCSFQLDLSLADNRDPHARVDYVAVQQQVFSPLPILPIANVKRILI	533
Qy	537	ANSRVGKGVMTFFGSLDFTSYLFIQEK	563
Db	534	SNTRVGDIEVTPFGNIHFKEKLSKKKK	53

RESULT 7

```

TV7M01_HAEDU
ID Q7VM01_HAEDU PRELIMINARY; PRT; 560 AA.
AC Q7VM01;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DE 07-FEB-2006, entry version 13.
DE Peptide transport, periplasmic protein Sapa.
GN Name=sapa; OrderedLocusNames=HD1230; ORFNames=HD_1230;
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=15000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungkur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi.".
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AE017143; AAP96068.1; -; Genomic DNA.
DR HSSP; P23847; 1DPE.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5; 1.
KW Complete proteome; -
SQ SEQUENCE 560 AA; 63809 MW; 9EE7FD98914355C8 CRC64;

```

Query Match 43.2%; Score 1267.5; DB 2; Length 560;
Best Local Similarity 44.8%; Pred. No. 1.2e-72;
Matches 251; Conservative 110; Mismatches 188; Indels 11; Gaps 4;

Qy	6	LRFSLGSLCISQSVLEQAPASVPTFLTEGLTYCTHAGSGFSPNPOTADAGTSMNVVTEQI	65
Db	9	LKFSFFFAVFCWISTAYSAPRVEKELSADSLIYCTSI SGLSFNPKQKADVGTNMNVTEQI	68
Qy	66	YNKLFEDKNHSATLLPMLQASYSISADGKEILLNLRHGKVFHQTPWTFPTDRFNAEDVVF	125
Db	69	YDKLFEDRTHRVFSLFETFSVSDGKEITLNLRAQVAFHCTPWFPTPTLRFNAEDVVF	128
Qy	126	SINRVGHNTYLTPL---AENATYVSNQPVRFVRLQKARVFYFDSIKLNEKISKVTAL	182
Db	129	SLNRMGTNVEELPALDFNEDSKEQFQGNQRYVAFKANLAIHYFPEVSVALKKKIASIP	185
Qy	183	SPYQVKIELFAPDSSILSHLSAQYAIIFSQEYAYQLSADNLAQLDTHPVGTGPYQVKDY	242
Db	189	NEYTVKIHVLVADPNVLAHLSAQYAVILSKEYALLMADEMLAQLDLLPVGTGQYVLSD	248
Qy	243	VYNQYVRLVRNENYVKEAKIEHIIVDLSDRSGLRVKFFNNECQIASYFPEVSQIGLLKN	302
Db	246	IQNEYVRLKFNPFVYGEKAKINNVDVDFSSTGRMAKYLQBCDQIVAOPEPSQRRISS	308
Qy	303	DDKHYYMQSDGMLNLAFLAFNDFKFLMRDHEIRAAISQSLNRARIHSYHNTATVANNI	362
Db	306	----YEIVSEGANLAFLAFNMKEQMQDIAFRRLQIANNRERLVAFLPFVGAEDVAV	364
Qy	365	IPESVASTVNTPEFEFDYHPKIAKNLAKDLNRLLIWINEQENYKVALPAPFKMAEMKWD	420
Db	365	LPSALFAOK--NPAAYPYKAPQPRAKNAKLDL---LIWFLEDSRVYLNHPLKMAEMRND	420

483 NLSNWCNEEFQDFMDRAITTSLSRAKAYNEAQELVLRPLPIPIANVKRILVANSRVK 542
 ||:||||:|:::| | | : | : : |||: : : | | : |
 481 NLANWCHQQFDEWLEIAKANQVPYVRNMIYKQTQALLEEQPLPLLLHAQRSFVNOKIK 540

Q32FZ3 SHIDS

```

ID Q3ZFZ3.SHIDS PRELIMINARY; PRT; 547 AA.
AC Q3ZFZ3;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE SapA.
GN Name=sapA; OrderedLocusNames=SDY_1638;
OS Shigella dysenteriae serotype 1 (strain Sd197).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300267;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong X., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery.";
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: CP000034; AB261762.1 -; Genomic DNA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
KW Complete proteome.
SQ SEQUENCE 547 AA; 61541 MW; 01AFB93EAC790FF3 CRC64;

```

Query Match 35.1%; Score 1030.5; DB 2; Length 547;
Best Local Similarity 37.7%; Pred. No. 1.8e-57;
Matches 210; Conservative 117; Mismatches 199; Indels 31; Gaps 9.

[illegible]

Db	345	RAYWYD-NEAKIT-EYNPAKSREQLSGLGENLTLLKLVAVPTRSQAWNPSPLKTAELIQ	402
Qy	422	DLAQAGVKVKVRAVTRPFLTAQLRNQSEYNDLISGLWAGNLDPDGMFRPILSCGTKNEL	481
Db	403	DLAQAGVKVKVIVPEGGEQEARLMDMS--HDLTSLSGWATSDNPDSFPLLSCAAHSQ	460
Qy	482	TNLNWCNEEFDQMDRAITTHLSLSRAKAYNEAQELVRLRPIPIANVKRILVANSRV	541
Db	461	TNLAHWCDFKFDVLRLKALSSQQLAARIEAYDEAQSILAQELPILPLASSILRLQAYRYDI	520
Qy	542	KGVMKTPFGSLDFTSLY	558
Db	521	KGVLSPFGNASFAGYV	537

Q83RL7.SHIFL
ID Q83RL7.SHIFL PRELIMINARY; PRT; 547 AA.
AC Q83RL7;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Peptide transport periplasmic protein.
GN Name=sapA; OrderedLocusNames=SF1299;
OS *Shigella flexneri*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=202772406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
RT through comparison with genomes of *Escherichia coli* K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

DR EMBL; AE005674; AAN42910.1; -; Genomic DNA.
DR GSSP; P23847; IDPE.
DR BioCyc; SFLE198214:AAN42910.1-MONOMER; -.
DR HGP; GO:0005215; E:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00495; SBP_bac_57.1.
DR PROSITE; PS01040; SBP_NUCLEOTIDE_5; 1.
KW Complete proteome.
SQ SEQUENCE 547 AA; 61548 MW; 8955A6DE587311D48 CRC64;

```

Query Match          35.0%; Score 1027.5; DB 2; Length 547;
Best Local Similarity 37.8%; Pred. No. 2.7e-7;
Matches 213; Conservative 112; Mismatches 194; Indels 45; Gaps 9;

Qy      9  LSLFLCISQSVLEQ--AAPSVPTF--LTENGLTYCTHAGSFNFNPQTADAGTSMNVVTEQ 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5  LSSLILVAGLVSGQAIAPESPAPHADIRDSFVYCVSGQVNTFNPKSSAGSLIVDTLAAQ 64

Qy     65  IYNKLFDIKNHSATLTPMLAQSYISADGKEILLMLRGVKFHTPPTPTPRDFNAEDVV 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     65  FYDRLLDQVDPYTYRLMPELASEVVLONGATYRFLHARDVFQKTDWTFPTRKNADVV 124

Qy    125  FSNIRVLGHNTYLPTEAANVTYSNPOYRVFHEQARKRVAFDFISKLNEIKSVTALP 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    125  FTFQIFRIFDN-----NFWIR-----NGSNMFYFDLSQFADNVKSVKRLDN 184

Qy    285  YQVCKELFADPSSILSHLAQYAIIFSQEYQLSDADNMLAQQLDTPVGTGPGVQVQDY 244
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    166  HTVEFLRAQDPASFLWHLHLATHYASVMSAEYARKLEKEDEKROOLDNPQVGTGPGY 225

```

Qy	245	NOYVRLVRNENYWKKEAKIEHIVDLSRSGRLKFFNNCEQIASYPEVSGILLKXND	304
Db	226	GOETRLQRHDDFWRGKELMPGVVDLGSQGTGRLSKLLTGCBCDVLAWPAASQLSILR-DD	284
Qy	305	KHYMMQSTDGMNLIAFNFDKFLMRDHEIRAAISQSLNRARIISHYHNATVANNIIP	364
Db	285	PRRLRLTEPGMNVYLAFAKTEPFLNPAVHRLALALINQRMQSIYYGTAETAASILP	344
Qy	365	EVSASVTNTPEFEFYDHKEI-----AKNK-----LADKNLLNLVINEQVYNPAFK	414
Db	345	RASWA-----YDNEAKITEYPAKRSREQKALGLENTI.LKILWPTRSQAWNPSPLK	395
Qy	415	MAEMIKWDLAQAGVHVRVRAVTPFLTAQLNRQSENYDILISGWLAGNLDPDGFMRPILS	474
Db	396	TAEIQA.DMAQGVGVVIVPEVGRFQARLMMS-----HDLTSGWATSDNPDSPFRPLLS	453
Qy	475	CTGKNELTNL.SNWCNEEFQDMRAITTHLSLRSKAKAYNEAQELVRLRELIPIANVKRI	534
Db	454	CAATHSQNTLAHWNPKPDSVLKRALSSQQLAARIEYAEQASILAQELPILPLASSRL	513
Qy	535	LVANSRGVGNVTPFSGLDSTLY 558	
Db	514	OAYRYDIGLGLV.SPFGNASFGVY 537	

RESULT 10

```

IDCW41 ECOL6
  DT 08CW41 ECOL6 PRELIMINARY; PRT: 547 AA.
AC 08CW41;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 21-FEB-2006, entry version 15.
DE Peptide transport periplasmic protein sapA.
GN Name=sapA; ORFNames=c_1771;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CPT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.;"
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL: AE014075; AA080237.1 -; Genomic DNA.
DR HSP: P23847; IDPE.
DR BioCyc: ECOL199310:C1771-MONOMER; -.
DR GO: GO:0005215; P:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR00914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5_1.
DR PROSITE: PS01040; SBP_BACTERIAL_5; 1.
KW Complete proteome.
SQ
SQ SEQUENCE 547 AA; 61466 MW; 48A9789B9C9C7947 CRC64;

```

```

Query Match          34.9%   Score 1025.5:  DB 2:  Length 547;
Best Local Similarity 37.8%   Prcd. No. 3.7e-57;
Matches 213;  Conservative 112;  Mismatches 194;  Indels 45;  Gaps 9;

Qy      9  LSFLLCISQSGVELQ--AAPSVPTF--LITENGLTYCTHAGSGFSINPQTADAGTSMNVVTEQ  64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5  LSSLLIVAGLVGSGQAIAAPESPPIADIRDGSGFVCYGVQNTVFNPSKASSGLIVDTLAAQ  64

Qy     65  IYNKFLDKIKNHSATLTPMLQAQSGKEILLNHHGVKFKTQFWFTTPRDFMAEDVV  124
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db     65  EYDLRLDQVPTYRIEMPELSESLVSDNGATVPHLRIDPEKFTQFTTPRDMVADV  124

```

```

Qy      125 FSINRVLGHNTYLTFLAEANVTYSNPQYRVFHEQARKVRFFYFDSIKLNEIKSVTALS 184
Db      125 FTFQIFDRN-----NFHWNV-----NGSNFFYFDSIQADNVKSVKRLDN 165

Qy      185 YQVKIELFAPDSSILSHLASQVAILFSQEVAYQLSADNLAQLDTHPGVGTGYQVKDVVY 244
Db      166 HTVEFLRAQPDASFLWHLATHYASVMSAEYAGLKEKEDRQEQLDROQPGVGTGYQLSEYRA 225

Qy      245 NQYVRLVRNENYWKKEAKIEHIIIVDLSTDRSGRLVKFFNNECQIASYPEVSGIGLKNDD 304
Db      226 GQYIRLQRHDDFWRGKPLMFQVQVVDLGSGGTGRLSKLLTGECVLAWPAASQLSILR-DD 284

Qy      305 KHYMQSTDGMNLAYLAFNFDKPLMRDHEIRAAISQSLNRARIHSIYHNTATVANNIIP 364
Db      285 PLRLTLRLPGMNVAYLAENTAKPELNNFAVRHALALAINNQLMQSIYGTAEATASLLE 344

Qy      365 EVSWASTVNTPEFEDYHPKI-----AKNK-----LADKNLLMLNVLVINEEQVNPAPFK 414
Db      345 RASWA-----YDNEAKITEYNPAKSREQKALGLENLTLKLVPTRSQAWNPSPLK 395

Qy      415 MAEMIKWDLAQAGVQVVKVRAVTRPFLTAQLRNQSENYDLILSGWLGNLDPDGMFRPLS 474
Db      396 TABELIQADMAQGVQVVIQVVEGRFQEARLMDMS--HDLTSLGWATSDNDPDSFFRPLLS 453

Qy      475 CGTKNELTNLSNWCNEEFQDMRAITTSLSLRKAYNEAQEVLRELPIPIANVKRI 534
Db      454 CAAHSQTNLAHWCDFKPFDSVLRKALSSQQLAARIEAYDKAQSLAQELFILPLASSLRL 513

Qy      535 LVANSRVKGVKMTFFGSLDPESTLY 558
Db      514 QAYRYDIKGLVLSFPGNASFAGVY 537

```

RESULT 11

SAPA_ECOLI

```

ID      SAPA_ECOLI      STANDARD;      PRT;      547 AA.
AC      Q47622; P77358;
DT      01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT      01-NOV-1996, sequence version 1.
DT      07-MAR-2006, entry version 38.
DE      Peptide transport periplasmic protein sapa precursor.
GN      Name=sapa; OrderedLocusNames=b1294;
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RC      NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP      STRAIN=K12 / FRAG5;
RA      Epstein W., Noelker E., Stumpe S., Tewes R., Schmid R., Bakker E.P.;
RL      Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474 (1997).
RN      [3]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=K12;
RX      MEDLINE=97251357; PubMed=9097039; DOI=10.1093/dnares/3.6.363;
RA      Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA      Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA      Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA      Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA      Sampei G., Seki Y., Sivasubramanian S., Tagami H., Takeda J.,
RA      Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT      "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT      corresponding to the 28.0-40.1 min region on the linkage map.";
RL      DNA Res. 3:363-377 (1996).
CC      -!- FUNCTION: Involved in a peptide intake transport system that plays
CC      a role in the resistance to antimicrobial peptides.

```

```

CC  -!- INTERACTION:
CC      POA6YS:dnaK; NbExp=1; IntAct=EBI-549564, EBI-542092;
CC  -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC  -!- SIMILARITY: Belongs to the bacterial solute-binding protein 5
CC      family.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; X97282; CAA65937.1; -; Genomic_DNA.
DR  EMBL; U00096; AAC74376.1; -; Genomic_DNA.
DR  EMBL; D90768; BAA14864.1; -; Genomic_DNA.
DR  EMBL; D90767; BAA14855.1; -; Genomic_DNA.
DR  PIR; A64878; A64878.
DR  HSSP; P23847; 1DPE.
DR  GenomeReviews; U00096_GR; b1294.
DR  EchoBASE; EB4155; -.
DR  EcoGene; EG20254; sapA.
DR  BioCyc; EcoCyc:SAPA-MONOMER; -.
DR  GO; GO:0005515; F:protein binding; IPI.
DR  InterPro; IPR000914; SBP_bac_5.
DR  Pfam; PF00496; SBP_bac_5; 1.
DR  PROSITE; PS01040; SBP_BACTERIAL_5; 1.
KW  Complete proteome; Peptide transport; Periplasmic; Protein transport;
KW  Signal; Transport.
FT  SIGNAL      1      21      Potential.
FT  CHAIN       22     547     Peptide transport periplasmic protein
FT                                     sapA.
FT                                     /FTId=PRO_0000031801.
FT  CONFLICT    34      35      SG -> RV (in Ref. 3).
SQ  SEQUENCE   547 AA;  61565 MW;  EB552BB3B8E102BF CRC64;

```

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006